Compact Electro-thermal Models of Semiconductor Devices with Multiple Heat Sources

C. Bohm, T. Hauck, E. B. Rudnyi, J. G. Korvink
Motorola GmbH, Schatzbogen7, 81829 München, Germany
IMTEK, Georges-Köhler-Alle 101-103, 79110 Freiburg, Germany
Christina.Bohm@motorola.com, Rudnyi@informatik.uni-freiburg.de

Abstract

The thermal management of semiconductor devices and systems becomes crucial as the power consumed by chips is increasing. For manufacturers it is important to enable the customer to simulate the thermal performance of semiconductor packages at different ambient conditions and arbitrary transient loading conditions. As simulation speed in this case is crucial, compact thermal models are of great importance.

In the present work, a new approach to generate compact thermal models of semiconductor packages on PCBs is presented. The main difference with conventional RC-like thermal networks is that the compact model is obtained through a formal model reduction procedure based on a solid mathematical background from the theory of approximation of large-scale dynamic system [3][4].

Model reduction starts with an accurate high-dimensional thermal model created with a finite element program like ANSYS. A low-dimensional model is obtained in such a way that the first moments of the transfer function are the same as in the original model. This is achieved by means of an iterative procedure (the Arnoldi algorithm) where during a single iteration one takes a previous reduced model of an order k and obtains the next reduced model of an order k+1.

For an automatic generation of the reduced order model the software mor4ansys has been developed. It reads system matrices of the original system from the binary ANSYS files and performs an Arnoldi algorithm in order to obtain a family of reduced models up to a maximum order specified by the user. The best order is chosen by a comparison of the results with the ANSYS simulation. Based on our experience, ANSYS models up to the order of 100 000 can be accurately represented by reduced model of an order 20-30 within the accuracy of several degrees.

A model of a semiconductor device with multiple heat sources is used in order to compare this method with a thermal RC network approach. Temperature response results to user defined loading conditions are also compared.

Theory

An accurate thermal model can be created by means of the finite element method. This can be routinely done by many software packages, for example ANSYS. After discretization in space, one obtains a system of ordinary differential equations

$$C\dot{\mathbf{T}} + K\mathbf{T} = \sum_i \mathbf{F}_i u_i \qquad (1)$$

where K and C are the system matrices, $\mathbf{T}$ is a vector of unknown temperatures at nodes made during the discretization, $\mathbf{F}_i$ is a load vector corresponding to the i-th heat source, and u_i is an input function, that is how the i-th heat power changes in time.

The Eq (1) can be directly transformed to the RC-like network to be used in the joint electro-thermal simulation [1] but the dimension of the state vector T is usually too high to use this opportunity in practice. As a result, one first should approximate the high-dimensional system (1) by a low-dimensional system [2], or in other words, derive a compact thermal model. In order to achieve this goal in the present work, we follow the approach presented in Ref [2] but employ the newest model reduction method [3][4]

An idea of model reduction is to find a low-dimensional subspace that describe the trajectory of the state vector with acceptable accuracy

$$\mathbf{T} = V\mathbf{z} + \boldsymbol{\varepsilon} \qquad (2)$$

that is, with the assumption that the error vector $\boldsymbol{\varepsilon}$ is small. The matrix V describes the transformation from the low-dimensional subspace of the generalized coordinates $\mathbf{z}$ to the original vector of unknown temperatures $\mathbf{T}$, its number of columns being equal the dimensionality of the low-dimensional subspace.

When the subspace V is found, one projects Eq (1) and obtains the low-dimensional system of ordinary differential equations

$$C_r\dot{\mathbf{z}} + K_r\mathbf{z} = \sum_i \mathbf{F}_{r,i} u_i \qquad (3)$$

where $K_r = V^T K V$, $C_r = V^T C V$, $\mathbf{F}_{r,i} = V^T \mathbf{F}$.

As one can integrate Eq (3) much faster than Eq (1), it can be used directly during joint electro-thermal simulation.

In order to find the transformation matrix V, we have used the moment matching method based on the block Arnoldi process [5]. A low-dimensional model is obtained in such a way that the first moments of the transfer function are the same as in the original model. An application of such a process to generate a compact thermal problem is presented in [6][7]. The main current practical limit of this approach is that it covers linear system of ordinary differential equations only.

Mor4ansys is a command-line tool built on top of the ANSYS-supplied library to read ANSYS binary files [8] and the TAUCS library for sparse linear algebra [9]. After a model is build and meshed in ANSYS, an ordinary differential equation (ODE) system of the first- or second-order is obtained. The current version of *mor4ansys* can handle ANSYS models up to 500 000 degrees of freedom. It is worthy of noting that the model reduction process employed in the present work is applicable not only to the thermal domain but rather to any linear problem in ANSYS: structural mechanics, acoustics including fluid-structure interaction, piezoelectric and electromagnetism [10][11].

The computational cost of model reduction is comparable with that of the stationary solution for Eq (1), the main computational effort being the Cholesky decomposition of the matrix K.

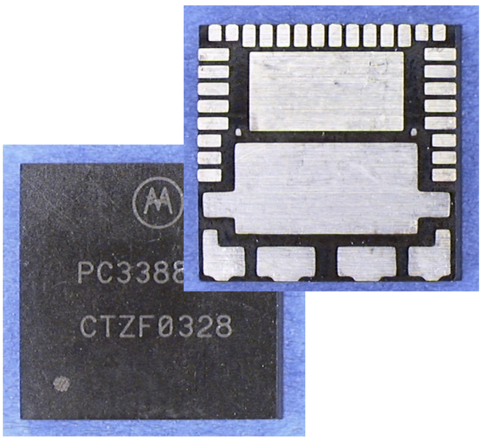

Figure 1: A PQFN Package

Finite Element Model of a Semiconductor Device

Transient thermal simulation for a Motorola semiconductor package mounted onto a multilayer PCB has been used to demonstrate the model reduction approach described above. The package is a PQFN with 10 mm x 10 mm outline dimensions used in automotive applications (see Figure 1). It is a leadless package with exposed die pad areas directly soldered to the printed circuit board. Two silicon chips are in this package, i.e., there are the two independently powered heat sources.

A detailed finite element model of this assembly has been created in ANSYS (ref. Figure 2). It has been linearized around the operation point. Unfortunately, this is a prerequisite in order to use the moment matching model reduction methods. The ambient conditions were set to natural convection at room temperature, implemented with a constant convection coefficient on the FE model surfaces.

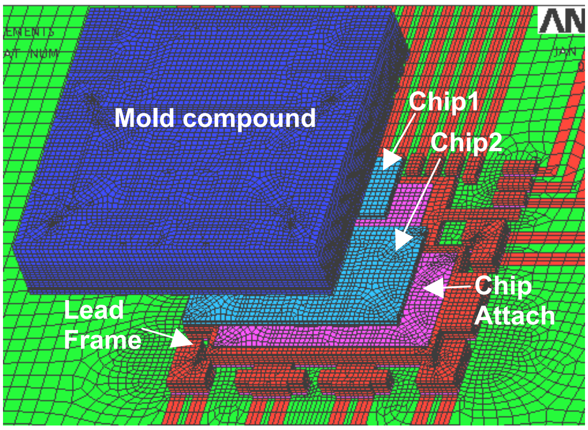

Figure 2: View of simulation model mounted onto PCB

Reduced Order Model – *mor4ansys*

The basic steps of how to generate the compact thermal model (3) from the ANSYS finite element model will be described in the following. A detailed manual for model reduction from ANSYS by means of the software *more4ansys* is available at IMTEK.

The first step is to create the finite element model in ANSYS and generate binary *.emat* files with system matrices and load vectors from Eq (1). The model should not exceed a 500000 degrees of freedom to be able to perform model reduction in reasonable time. Unfortunately, ANSYS can save only a single load vector in the *.emat* file. As a result, for each of the n heat sources that are independently powered, the extraction of n *.emat* files is required. Then, these *.emat* files, as well as the lists for the Dirichlet boundary conditions and output degrees of freedoms have been uploaded to the IMTEK server via ftp. IMTEK has created a Web-based interface where the user starts the automated procedure of model reduction from the uploaded files. For our example, an order of the reduced matrices has been selected to be equal to 25. The results, i.e., the reduced matrices K_r and C_r, load vectors $\mathbf{F}_{r,i}$ and the transformation matrix V, have been downloaded back in an ASCII file.

RC Network Model

The generation of conventional thermal RC networks for transient thermal simulation is described elsewhere [12]. This procedure to create a thermal RC network that is appropriate to represent the assembly in Fig. 2 is semi-automatic. This means that first the user has to create the network topology, which depends on the number of heat sources and complexity of the structure. Then, the parameters of the RC network are extracted automatically from fitting the parameters of the transfer functions to the step response of the structure that are either measured or, in our case, calculated from the finite element ANSYS model.

For our example, 3 and 4 rung ladders of RC cells considering also the interaction between the heat sources have been created. The system matrices are created from the network parameters, the matrix dimension being equal to 14. At source nodes, loading conditions are defined and temperatures are calculated, while results at inner nodes for this RC network have no physical meaning. The resulting set of equations can be processed in Mathematica for thermal simulation.

$$\begin{bmatrix}[C_{aa}] & [C_{ai}]\\ [C_{ia}] & [C_{ii}]\end{bmatrix}\begin{Bmatrix}\{\dot{T}_a\}\\ \{\dot{T}_i\}\end{Bmatrix} + \begin{bmatrix}[K_{aa}] & [K_{ai}]\\ [K_{ia}] & [K_{ii}]\end{bmatrix}\begin{Bmatrix}\{T_a\}\\ \{T_i\}\end{Bmatrix} = \begin{Bmatrix}\{P_a\}\\ \{0\}\end{Bmatrix}$$

(*a indicates source nodes, i indicates inner nodes*) (4)

Model Response Comparison at Step- and Pulse Stimulus Functions

In both cases of compact thermal models either by means of matrix reduction (reduced order model or RO model) or by thermal RC networks, the two matrices (K and C) are available that can be used for transient thermal simulation. In the case of model reduction, the additional transformation matrix V (ref. Eq. 2) is to be considered to convert the generalized coordinates to temperatures (see Eq 3). The time integration for transient thermal simulation with the reduced models was implemented by means of the finite difference method in Mathematica.

Unit step responses were calculated to compare simulation results of the FE, RO and RC network model. For the simulation results shown in Fig 3, Chip2 is powered, while Chip1 is passively heated. As one can see, there is a good agreement among all three simulation models. Over the considered time range, the difference of both, RC and RO models, to the FE results stays below 1°C.

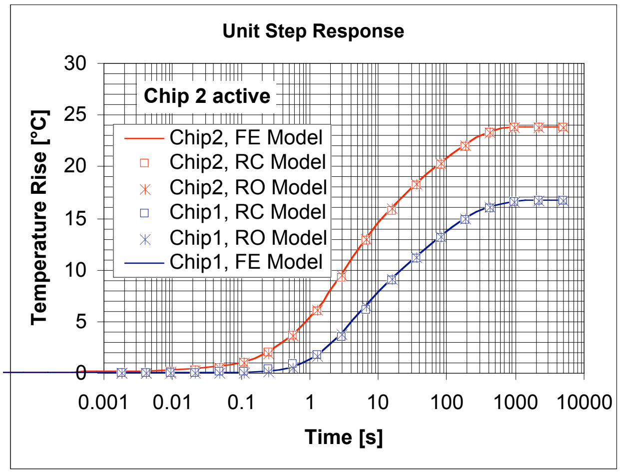

Figure 3: Unit step response comparison

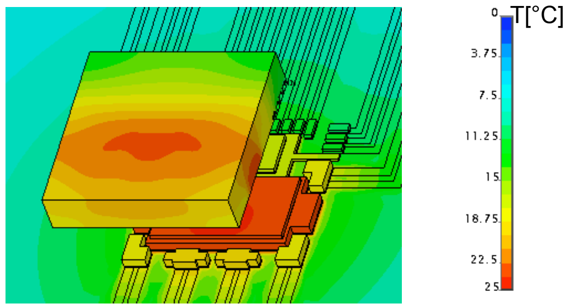

Figure 4: View of finite element temperature field solution

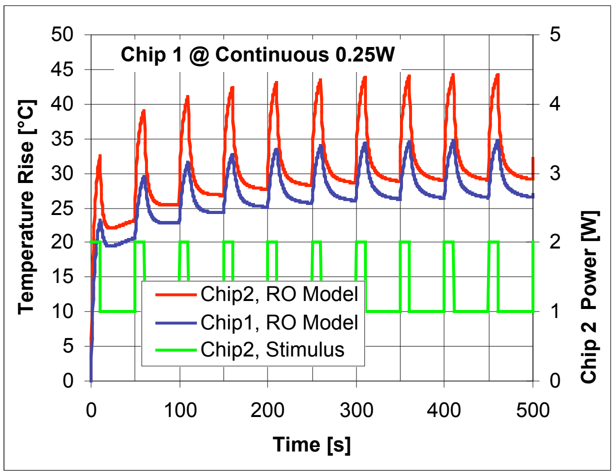

Figure 5: Reduced order model response to arbitrary loading conditions

For arbitrary transient loading conditions, thermal simulation with RO and RC model lead to a significant reduction in the simulation time. Figure 5 shows temperature step responses for both heat sources. Chip1 is continuously powered with 0.25W while pulse loads are applied to Chip2. The temperature response in both chips is calculated within seconds, while a finite element simulation takes significantly larger amount of calculation time and work space.

Discussion

The basic steps to create compact thermal models from the finite element model were demonstrated on the example of a semiconductor device assembly. The two methods of creating reduced order models were compared: an automated procedure provided by IMTEK that reduces the finite element matrices by means of the moment matching method via the Arnoldi algorithm vs. thermal RC networks.

The results from both methods are lower order system matrices that can be used for fast transient thermal simulations. A comparison of unit step response solutions shows that both reduced order models possess small differences only as compared to the finite element results. The main benefits of reduced order simulation is significant saving in computational time and work space.

As usual each method has advantages and disadvantages. The advantage of the method developed at IMTEK is that it is applied directly to system matrices, and as such it is completely automatic and very efficient computationally. The computational cost of the model reduction is comparable with that of the stationary solution in ANSYS. A user should just develop his/her model in ANSYS and then just run the software mor4ansys, as the method is based on a formal mathematical procedure. Another advantage is that the method is quite general in nature and can be applied to other ANSYS problems, for example, mechanical.

On the other hand, the RC network method uses as input information the transient simulation result. For this reason, it can be applied not only to FEM models but also to experimentally measured responses. This can be an advantage in the case when there are uncertainties in the material thermophysical properties as well as in the boundary conditions. However, this method requires a step of choosing the right network topology that it is quite difficult to make completely automatic.

Acknowledgements

We acknowledge partial funding by the German Ministry of Research BMBF. IMTEK also acknowledges the partial funding by the DFG project MST-Compact (KO-1883/6) and by an operating grant of the Freiburg university.

References

[1] J. T. Hsu and L. VuQuoc, "A rational formulation of thermal circuit models for electrothermal simulation. 1. Finite element method," Ieee Transactions on Circuits and Systems I-Fundamental Theory and Applications, vol. 43, pp. 721-732, 1996.

[2] J. T. Hsu and L. VuQuoc, "A rational formulation of thermal circuit models for electrothermal simulation. 2. Model reduction techniques," Ieee Transactions on Circuits and Systems I-Fundamental Theory and Applications, vol. 43, pp. 733-744, 1996.

[3] A. C. Antoulas, D. C. Sorensen, Approximation of large-scale dynamical systems: An overview, Technical Report, Rice University, Houston, 2001, http://www-ece.rice.edu/~aca/mtns00.pdf.

[4] E. B. Rudnyi, J. G. Korvink, Automatic Model Reduction for Transient Simulation of MEMS-based Devices, Sensors Update, 2002, 11, 3-33.

[5] R. W. Freund, “Krylov-subspace methods for reduced order modeling in circuit simulation”, Journal of Computational and Applied Mathematics 123, 395–421, 2000.

[6] T. Bechtold, E. B. Rudnyi, and J. G. Korvink, "Automatic generation of compact electro-thermal models for semiconductor devices," Ieice Transactions on Electronics, vol. E86C, pp. 459-465, 2003.

[7] L. Codecasa, D. D'Amore, and P. Maffezzoni, "An Arnoldi based thermal network reduction method for electro- thermal analysis," Ieee Transactions on Components and Packaging Technologies, vol. 26, pp. 186-192, 2003.

[8] Guide to interfacing with Ansys, ANSYS corporation.

[9] S. Toledo, D. Chen, V. Rotkin, TAUCS – A library of sparse linear solvers, http://www.tau.ac.il/~stoledo/taucs/.

[10] E. B. Rudnyi, J. Lienemann, A. Greiner, and J. G. Korvink, mor4ansys: Generating Compact Models Directly from ANSYS Models, accepted for publication in the Technical Proceedings of the 2004 Nanotechnology Conference and Trade Show, Nanotech 2004, March 7-11, 2004, Boston, Massachusetts, USA.

[11] J. Lienemann, D. Billger, E. B. Rudnyi, A. Greiner, and J. G. Korvink, MEMS Compact Modeling Meets Model Order Reduction: Examples of the Application of Arnoldi Methods to Microsystem Devices, accepted for publication in the Technical Proceedings of the 2004 Nanotechnology Conference and Trade Show, Nanotech 2004, March 7-11, 2004, Boston, Massachusetts, USA

[12] T.Hauck, Thermal RC-network approach to analyze multichip power packages, Semitherm, 2000